1 – 25 (Previously Canceled)

26. A method of populating a natural language speech lattice with semantically variant questions, the method comprising the steps of:

(a) providing a target set of questions associated with a content of a task domain that is to be supported by the natural language speech lattice;

(a)' receiving a first user question from said target set of questions;

(b) dividing the user question into a plurality of words corresponding to the user question;

(c) determining synonyms for [each] selected words in said plurality of words;

(d) formulating a [random] synonym set of questions related to said user question based on said synonyms;

(e) performing semantic decoding on said [random] synonym set of questions, to identify a disambiguated set of questions;

(f) storing said set of disambiguated questions in a speech recognition lattice;

wherein said set of disambiguated questions correspond to semantic variants of said target set of questions that can be [posed to] recognized by a natural language speech engine used for the task domain.

27. (Currently amended) The method of claim [25] 26, wherein steps (a) through (f) are performed by a software program embodied in a machine readable media.

28. (Currently amended) The method of claim [25] 26, wherein the method is performed by a natural language engine operating on an online server connected to the Internet.

29. (New) The method of claim 26, further including a step: receiving a new user question and correlating the same to one or more of said set of disambiguated questions.

30. (New) The method of claim 29 further including a step: providing one or more answers to said new user questions.

31. (New) The method of claim 30, wherein statistical processing is also performed, including measuring an overlap of phrases, to identify said one or more answers.

32. (New) The method of claim 26, wherein said lattice is incorporated into an N-gram statistical language model.

33. (New) The method of claim 26, wherein said synonyms are derived from a programmatic lexical dictionary, including WordNet.

34. (New) The method of claim 26 wherein said semantic decoding is performed by computing a semantic distance between like parts of speech in said synonymic set of questions and said user question.

35. (New) The method of claim 26 wherein said selected words also include word phrases, and said synonyms include phrasal synonyms.

36. (New) The method of claim 26 wherein said target set of questions are automatically derived from transcriptions of user utterances.

37 (New) The method of claim 26 wherein said target set of questions are derived from said content.

38. (New) A method of populating a natural language speech lattice with semantically variant questions, the method comprising the steps of:

(a) defining a first set of topic questions associated with content of a task domain to be supported by the natural language speech lattice;

wherein said first set of topic questions also each have at least one corresponding topic answer which can be provided by a speech based natural language system for the task domain in response to a speech based query;

(b) for each topic question;

i) dividing the topic question into a plurality of words corresponding to the topic question;

ii) determining semantically related words for each word in said plurality of words;

iii) formulating a semantic set of questions related to the topic question based on said semantically related words;

iv)   performing semantic decoding on said semantic set of questions for the topic question, to identify a disambiguated set of questions appropriate for the topic question;

v)   storing said set of disambiguated questions for the topic question in a speech recognition lattice;

(C) repeating step for each topic question in said first set of topic questions;

wherein said set of disambiguated questions for said first set of topic questions correspond to semantic variants of speech based queries that are supported by said speech based natural language system for the task domain.

39. (New) The method of claim 38 wherein said semantically related words include at least one of a synonym, hyponym and/or hypernym.

40. (New) A method of generating a statistical language model (SLM) grammar for a task domain which includes semantically variant words and phrases, the method comprising the steps of:

(a) providing a set of content words which can be associated with user questions in the task domain;

(b) determining semantic variants for each word in said set of content words;

wherein said semantic variants include at least synonyms;

(d) forming a semantic set of questions related to said user questions based on said synonyms;

(e) performing semantic decoding on said semantic set of questions, to identify a disambiguated set of questions;

(f) configuring n-gram probabilities for words and phrases in said SLM grammar based on said set of disambiguated questions;

wherein said SLM grammar is configured to recognize semantic variants of questions posed to a natural language speech recognition engine.

41. (New) The method of claim 40, wherein said set of content words are extracted automatically from files containing text transcriptions of user utterances.

42. (New) The method of claim 40 further including a step: adjusting n-gram probabilities of the SLM grammar for the task domain based on observation count data from a second larger SLM grammar with a different set of words and phrases.

43. (New) The method of claim 40, wherein said n-gram probabilities are based on bi-grams.

44. (New) The method of claim 40 further including embedding steps (a) through (f) into a software data preparation tool.